Journal of Applied Sciences 10 (1): 59-64, 2010
ISSN 1812-5654

A Novel Embedding Method to Increase Capacity and Robustness of Low-bit Encoding Audio Steganography Technique Using Noise Gate Software Logic Algorithm

Mohamed A. Ahmed, Miss Laiha Mat Kiah, B.B. Zaidan and A.A. Zaidan
Faculty of Computer Science and Information Technology,
University of Malaysia, 50603, Kuala Lumpur, Malaysia

Abstract: In this study, we developed a novel method that is able to shift the limit for transparent data hiding in audio from the fourth LSB layer to the eighth LSB layer, thus the method has improved the capacity and robustness of data hiding in the audio file using a two steps approach. In the first step, noise gate software logic used to obtain a desired signal for embedding the secret message of the input host audio signal (carrier). In the second step, standard 8th LSB layer embedding has been done for this desired signal. Audio quality evaluation has used to evaluate our purposed method in two ways. First, objective test showed the algorithm succeeds in this task, while increasing SNR values of our algorithm comparing to SNR values obtained by standard LSB embedding in the 8th bits LSB layer and the comparison of the histogram audio excerpts has proved this also. Second, subjective listening test proved that high perceptual transparency is accomplished even if eight LSB of host audio signal are used for data hiding.

Key words: Data hidden, audio steganography, SNR, steganography, noise gate, least significant bits

INTRODUCTION

Steganography is the art and science of writing hidden messages in such a way that no-one can realizes there is a hidden message in data (e.g., images file, documents file, sounds file etc.) except the sender and intended recipient, the word steganography is of Greek origin and means covered, or hidden writing. Steg analysis concentrates on the art and science of searching and or destroying secret messages that have been produced using any of the various steganography techniques. Cryptography obscures the meaning of a message, but it does not conceal the fact that there is a message, cryptography is the art of secret writing, which is intended to make a message unreadable by a third party but does not hide the message in a communication. Cryptanalysis is analogous to steganalysis but it applied with cryptography rather than steganography (Artz, 2001).

Although, steganography is separate and distinct from cryptography, but there are many analogies between the two and some authors categorize steganography as a form of cryptography since hidden communication is a form of secret writing (Artz, 2001).

A Watermark is a recognizable image or pattern in paper that appears as various shades of lightness/darkness when viewed by transmitted light (or when viewed by reflected light, atop a dark background), caused by thickness variations in the study. Digital watermarking is the process of embedding information into a digital signal, the signal may be audio, pictures or video, for example. If the signal is copied, then the information is also carried in the copy. Watermarking systems can be characterized by a number of properties (Stefan and Fabin, 2000). The relative importance of each property depends on the requirements of the system application. The property that has been associated with our research is robustness; robustness refers to the ability to detect the watermark after common signal processing operations. Audio watermarking needs to be robust to temporal filtering, A/D conversion, time scaling, etc. Not all applications of watermarking require all the forms of robustness, this depends on the nature of application of watermarking system (Alsalami and Al-Akaidi, 2003). Steganography and digital watermarking are the same but the purpose of last one is use to copyright protection systems, which are intended to prevent or deter unauthorized copying of digital media. Steganography is an application of digital watermarking, where two parties communicate a secret message embedded in the digital signal.

Classical Steganography often used methods of completely obscuring the message so it was unnoticeable to those who didn't know the specific covert method it

Corresponding Author: Mohamed. A. Ahmed, Faculty of Computer Science and Information Technology, University of Malaysia, 50603, Kuala Lumpur, Malaysia Tel: +060172452457

was using for example Invisible inks, which was able to write a confidential letter with any other non-value-confidential and usually write between lines (Anderson and Petitcolas, 1998). Modern Steganography refers to hide information in digital picture, audio or text files etc., each one of this digitals data has a many techniques can use with it , for example in digital image the JPHide/JPSeek uses the coefficients in a JPEG to hide information, this method alter the image. In digital audio file several packages also exist for hiding data in audio files, such as MP3Stego not only effectively hides arbitrary information, but also claims to be a partly robust method of watermarking MP3 audio files (Noto, 2001). The windows wave format lets users hide data using Steghide, it alters the Least Significant Bits (LSB) of data in the carrier medium (Artz, 2001).

Basically, all steganography techniques have to satisfy two basic requirements. The first requirement is perceptual transparency or noticeable perceptual distortion i.e., cover object or carrier (object not containing any additional data) and stego object (object containing secret message) must be perceptually indiscernible (Anderson and Petitcolas, 1998), the cover object and stego object have the same media type like image or video or audio file, in this research the media type will be digital audio so the research problem will be in audio steganography. The second requirement is high data rate of the embedded data. The main challenge in digital audio watermarking and steganography is that if the perceptual transparency parameter is fixed, the design of a watermark system cannot obtain high robustness and a high watermark data rate at the same time (Cvejic and Seppanen, 2004).

All the stego-applications, besides requiring a high bit rate of the embedded data have needed of techniques or algorithms that detect and decode hidden bits, hiding data in audio can be done a number of ways like: Phase coding, Spread spectrum, Echo data hiding, Patchwork coding, Low-bit encoding (LSB-based) (Bender *et al*., 1996). However, Chang and Moskowitz (1997) were analyzed these techniques for information hiding in digital audio and among them the Low-bit encoding technique has the highest watermark channel bit rate but with low robust.

Phase coding works by substituting the phase of an initial audio segment with a reference phase that represents the hidden data. The phase of subsequent segments is adjusted in order to preserve the relative phase between segments (Bender *et al*., 1996), so the basic idea is to split the original audio stream into blocks and embed the whole watermark data sequence into the phase spectrum of the first block. One drawback of the phase coding method is a considerably low payload because only the first block has used for watermark embedding (Cvejic and Seppanen, 2004).

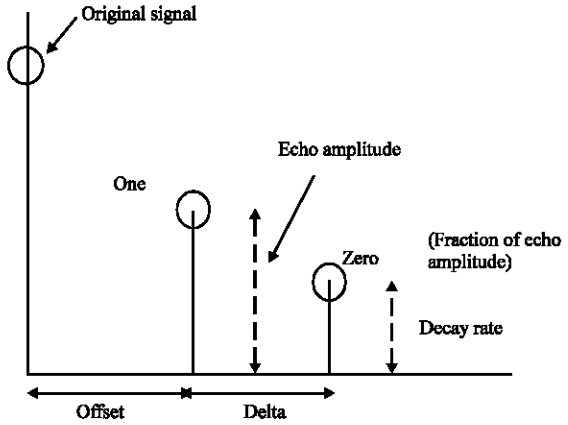

Fig. 1: Echo data hiding adjustable parameters

Spread Spectrum (SS) is technique designed to encode a stream of information by spreading the encoded data across as much of the frequency spectrum as possible. This allows the signal reception, even if there is interference on some frequencies (Bender *et al*., 1996). There are two variations on spread spectrum communication and its can be used in information hiding and audio steganography such Direct Sequence Spread Spectrum (DSSS) coding and Frequency Hopping Spread Spectrum (FHSS) coding (Cvejic and Seppanen, 2007; Stefan and Fabin, 2000).

Echo data hiding embeds data into a host audio signal by introducing an echo. The data are hidden by varying three parameters of the echo: initial amplitude, decay rate and offset. As the offset (or delay) between the original and the echo decreases, the two signals blend. At a certain point, the human ear cannot distinguish between the two signals. The coder uses two delay times, one to represent a binary one (offset) and another to represent a binary zero (offset + delta) (Fig. 1).

MATERIALS AND METHODS

Low-bit encoding (LSB-based): It was the earliest techniques studied in the information hiding and watermarking area of digital audio (Cvejic and Seppanen, 2002; Yeh and Kuo, 1999) as well as other media types (Lee and Chen, 2000) the main advantage of this technique is a very high watermark channel capacity; the use of only one least significant bit of the host audio sample gives the capacity of 44.1 kbps if a mono audio signal, sampled at 44.1 kHz, was used. The obvious disadvantage is the method's extremely low robustness,

due to the fact that the random changes of the LSB destroy the coded watermark (Mobasseri, 1998). Ideally, the channel capacity is 1 kbps per 1 kHz, e.g., in a noiseless channel, the bit rate will be 8 kbps in an 8 kHz sampled sequence and 44 kbps in a 44 kHz sampled sequence and so on (Noto, 2001; Bender *et al*., 1996).

As the number of used LSB during Low-bit encoding increases or, equivalently, depth of the modified LSB layer becomes larger, probability of making the embedded message statistically detectable increases and perceptual transparency of stego objects will decrease. Therefore, there is a limit for the depth of the used LSB layer in each sample of host audio that can be used for data hiding. In the literature survey for this technique, audio quality evaluation showed that, in average, the maximum LSB depth that can be used for Low-bit encoding watermarking without causing noticeable perceptual distortion is the fourth LSB layer when 16 bits per sample audio sequences are without any methods (embedded in the bits of all samples directly), but with method or algorithm succeeds in increasing the depth of the embedding layer from 4th to 6th LSB layer without affecting the perceptual transparency of the watermarked audio signal Cvejic and Seppanen (2004, 2005).

Proposed low-Bit encoding method: We developed a novel method that is able to shift the limit for transparent data hiding in audio from the fourth LSB layer to the eighth LSB layer, using a two steps approach. In the first step, noise gate software logic used to obtain a desired signal for embedding the secret message of the input host audio signal (carrier). In the second step, standard 8th LSB layer embedding has been done for this desired signal. The standard low -bit encoding method simply replaces the original host audio bit in the ith layer (i=1,...,16) with the bit from the watermark bit stream.

The key idea of our proposed algorithm is to avoid embedding in a silent periods of host audio signal or any point that near from this silent periods (regardless the style of music genres) due to sensitivity of human auditory system (HAS) for these periods so it will affect the perceptual transparency or noticeable perceptual distortion in a certain manner.

Noise gate software logic (Davis and Jones, 1989) has used to control embedding in a silent periods of host audio signal (carrier) or any point that near from these silent periods. It allows an embedding to pass through only when it is equal or above a predefined threshold: the gate is 'open' and the output will be the embedding secret message with the carrier otherwise if the signal falls below the threshold: the gate is 'closed' and only the carrier will be allowed to pass without carrying any valid message (additional data).

In the encoding operation (embedding) (Fig. 2), the comparator (Noise Gate Software Logic) has used to open/close or enable/disable the buffers. If the value of input host audio signal is equal or greater than value of threshold, the output of comparator will be (on or 1), this will enable or open buffer (A) and disable or close buffer (B), so the final output will be the host audio signal with secret message (standard 8th LSB layer embedding).

Otherwise, if the value of input host audio signal is less than value of threshold, the output of comparator will be (off or 0), this will enable or open buffer (B) and disable or close buffer (A), so the final output will be same of input or the host audio signal only without any secret message or without embedding.

In the decoding operation (Extraction) see (Fig. 3), the comparator (Noise Gate Software Logic) has used to enable or disable the (Extractor). If the value of input stego audio signal is equal or greater than value of threshold, the output of comparator will be (on or 1), this will enable the (extractor), so the final output will be the secret message (standard 8th LSB layer extracting).

Otherwise, if the value of input stego audio signal is less than value of threshold, the output of comparator will be (off or 0), this will disable the (extractor), so the final output will be nothing or no output.

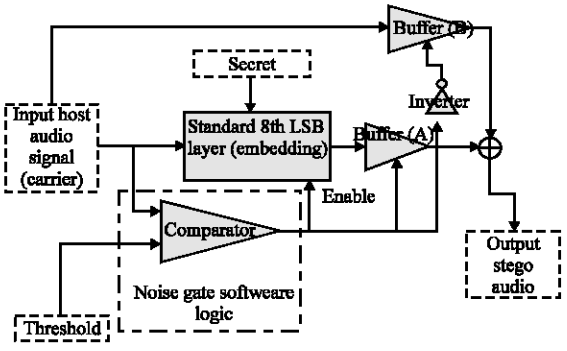

Fig. 2: Embedding audio steganography (encoding)

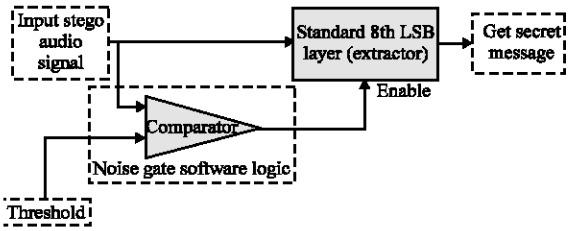

Fig. 3: Extraction audio steganography (decoding)

Therefore, we expect that, using the proposed two-step algorithm, we can increase the depth of watermark embedding further than the 4th LSB layer and accordingly increase algorithm's robustness towards noise addition.

RESULTS AND DISCUSSION

Audio quality evaluation (Arnold, 2000; Lin and Abdulla, 2008) has used to evaluate our purposed method in two ways. First, subjective listening tests that concern with human's sensitivity evaluation, so we will use Subjective Difference Grade (SDG) here. Second, objective evaluation tests that concern with statistical results of machine calculations, so we will use (SNR) and (histogram) here.

Our proposed LSB watermarking algorithm was tested on 10 audio files sequence from different music styles (speech {1}, rock {2, 3}, techno {4, 5, 6}, jazz {7, 8}, country {9, 10}). The audio excerpts were selected so that they represent a broad range of music genres, i.e., audio clips with different dynamic and spectral characteristics. All music pieces have been watermarked using the our proposed algorithm (embedding layer was 8th LSB) and standard LSB watermarking algorithm (also embedding layer was 8th LSB). Clips were 44.1 kHz sampled mono audio files, represented by 16 bits per sample. Duration of the samples ranged from 10 to 20 sec. As defined in (Bassia and Nikolaidis, 2001) signal to noise ratio in a time domain for the embedded watermark is computed as:

$$SNR = 10 \times \log_{10} \frac{\sum_1^n x^2(n)}{\sum_1^n [x(n) - y(n)]^2}$$

where, n represents the total number of samples per audio file, x(n) represent a sample of input audio sequence (before embedding) and y(n) stands for a sample of audio with modified LSB (after embedding).

In SNR of watermarked audio signals, a higher SNR reveals that the watermarked audio signals are much closer to the host signal (Cvejic and Seppanen, 2007), so it can be seen from Fig. 4, that our algorithm outperforms for SNR values from standard LSB insertion algorithm (threshold = 0) to our novel LSB insertion algorithm (threshold =256, 512, 768, 1024; so on).

Subjective quality evaluation of the watermarking method was performed by listening tests involving ten persons. Four of them had basic or medium level music education or are active musicians. In the first part of the test, participants listened to the original and the watermarked audio sequences and were asked to report dissimilarities between the two signals, using a 5-point impairment scale (SDG): (5: Imperceptible, 4: perceptible but not annoying, 3: slightly annoying, 2: annoying 1: Very annoying). Table 1 presents results of the first test (SDG), with the average Mean Opinion Score (MOS) for four of the 10 tested audio excerpts.

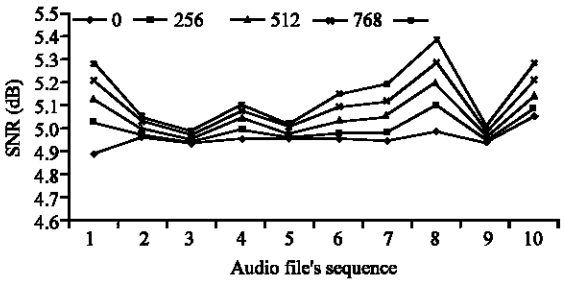

Fig. 4: SNR values for tested audio files sequence

Table 1: Shows the results of the first test (SDG), with the average Mean Opinion Score (MOS) for four of the ten tested audio excerpts

Audio steganography	Speech	Jazz	Rock	Country
Standard method (threshold = 0)	4.4	4.5	4.7	3.9
Our method (threshold = 256)	4.9	4.7	4.9	4.5
Our method (threshold = 512)	5.0	4.9	5.0	4.9
Our method (threshold = 768)	5.0	5.0	5.0	5.0
Our method (threshold = 1024)	5.0	5.0	5.0	5.0

Table 2: Shows the blind audio watermark test for four of the ten tested audio excerpts

Audio steganography	Speech	Jazz	Rock	Country
	(%)			
Standard method (threshold = 0)	43	45	57	40
Our method (threshold = 256)	45	48	53	48
Our method (threshold = 512)	49	49	50	49
Our method (threshold = 768)	50	50	50	50
Our method (threshold = 1024)	50	50	50	50

In the second part, test participants were repeatedly presented without watermarked and with watermarked audio clips in random order and were asked to determine which one is the watermarked one (blind audio watermarking test). Experimental results are shown in a Table 2, where values near to 50% show that the two audio clips (original audio sequence and watermarked audio signal) cannot be discriminated by people that participated in the listening tests.

Results of subjective tests showed that perceptual quality of watermarked audio, if embedding is done using the novel algorithm, is higher in comparison to standard LSB embedding method. Discrimination values and mean opinion scores in the case of standard and proposed algorithm embedding are obviously different. This confirms that described algorithm succeeds in increasing the depth of the embedding layer from 4th to 8th LSB layer regarding to the perceptual transparency of the watermarked audio signal. Therefore, a significant improvement in robustness against signal processing manipulation can be obtained.

In order to compare the between our proposed algorithm and the standard one, last objective test has

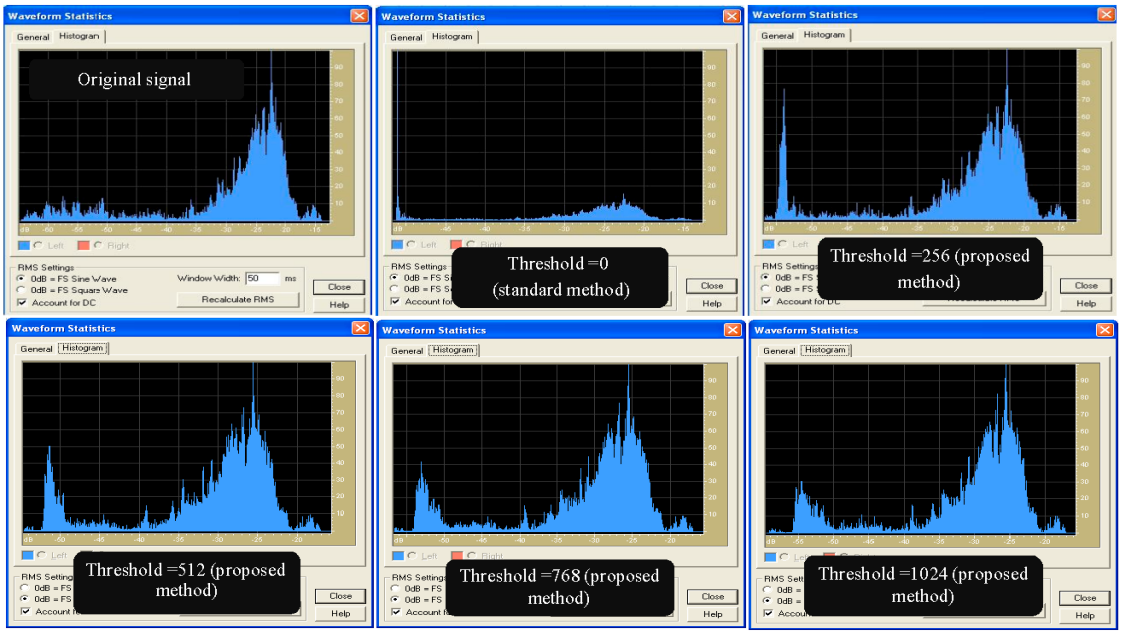

Fig. 5: Illustrated the histogram of certain audio file chosen randomly from audio files sequence

been done using histogram statistics figures for audio file before and after embedding operation, Fig. 5 has illustrated the histogram of certain audio file chosen randomly from audio files sequence and you can compare the histogram of original audio signal (without embedding) with the other ones (with embedding), standard LSB insertion algorithm (threshold = 0) to our proposed insertion algorithm (threshold = 256, 512, 768 and 1024). However, if you look at figures you can sense or see when we have applied our proposed algorithm the shape of signal has returned or approximately returned to the original shape, this also confirms that described algorithm succeeds in increasing or improving the capacity and robustness of data hiding in the audio file.

CONCLUSIONS

We presented a novel algorithm for low-bit encoding audio steganography technique. The key idea of our proposed algorithm is to avoid embedding in a silent periods of host audio signal or any point that near from this silent periods (regardless the style of music genres) due to sensitivity of Human Auditory System (HAS) for these periods so it will affect the perceptual transparency or noticeable perceptual distortion in a certain manner. Noise gate software logic has used to control embedding in a silent periods of host audio signal (carrier) or any point that near from these silent periods. Audio quality evaluation showed that described algorithm succeeds in increasing the depth of the embedding layer from 4th to 8th LSB layer without affecting the perceptual transparency of the watermarked audio signal, thus the method has improved the capacity and robustness of data hiding in the audio file. Objective test showed the algorithm succeeds in this task, while increasing SNR values of our algorithm comparing to SNR values obtained by standard LSB embedding in the 8th bits LSB layer and the comparison of the histogram audio excerpts has proved this also. Subjective listening test proved that high perceptual transparency is accomplished even if eight LSB of host audio signal are used for data hiding.

ACKNOWLEDGMENTS

Our sincere thanks to all researchers who have contribute to this project. Also, we would like to acknowledge and thanks the researchers in UM for their support to some idea presented here were classified by discussion with baraa albaker.

REFERENCES

Alsalami, M.A.T. and M.M. Al-Akaidi, 2003. Digital Audio Watermarking Survey. De Montfort University, UK., pp: 1-14.

Anderson, R.J. and F.A.P. Petitcolas, 1998. On the limits of steganography. IEEE J. Selected Areas Commun., 16: 474-481.

Arnold, M., 2000. Audio watermarking Features applications and algorithms. Proc. IEEE Int. Conf. Multimedia Expo, 2: 1013-1016.

Artz, D., 2001. Digital steganography: Hiding data within data. IEEE Internet Comput., 5: 75-80.

Bassia, P., I. Pitas and N. Nikolaidis, 2001. Robust audio watermarking in the time domain. IEEE Trans. Multimedia, 3: 232-241.

Bender, W., D. Gruhl, N. Morimoto and A. Lu, 1996. Techniques for data hiding. IBM Syst. J., 35: 313-336.

Chang, L. and I.S. Moskowitz, 1997. Critical analysis of security in voice hiding techniques. Proceedings of the 1st International Conference on Information and Communications Security Beijing, China, Nov. 11-14, Springer, Berlin, Heidelberg, pp: 203-216.

Cvejic, N. and T. Seppanen, 2002. Increasing the capacity of LSB-based audio steganography. Proceedings of the 5th IEEE Workshop on Multimedia Signal Processing, Dec. 9-11, St. Thomas, VI, pp: 336-338.

Cvejic, N. and T. Seppanen, 2004. Reduced distortion bit-modification for LSB audio steganography. Proc. 7th Int. Conf. Signal Process., 3: 2318-2321.

Cvejic, N. and T. Seppanen, 2005. Increasing Robustness of LSB audio steganography by reduced distortion LSB coding. J. Univ. Comput. Sci., 11: 56-65.

Cvejic, N. and T. Seppanen, 2007. Digital Audio Watermarking Techniques and Technologies Applications and Benchmarks. IGI Global, USA. UK., ISBN-10: 1599045133, pp: 328.

Davis, G. and R. Jones, 1989. The Sound Reinforcement Handbook. 2nd Edn., Hal Leonard Corporation, Milwaukee.

Lee, Y.K. and L.H. Chen, 2000. High capacity image steganographic model. IEEE Proc. Vision Image Signal Process., 147: 288-294.

Lin, Y. and W.H. Abdulla, 2008. Perceptual evaluation of audio watermarking using objective quality measures. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, March 31-April 4, Las Vegas, Nevada, USA., pp: 1745-1748.

Mobasseri, B.G., 1998. Direct sequence watermarking of digital video using m-frames. Proc. IEEE Int. Conf. Image Process., 2: 399-403.

Noto, M., 2001. MP3Stego: Hiding text in MP3 files. http://www.sans.org/reading_room/whitepapers/steganography/mp3stego_hiding_text_in_mp3_files_550.

Stefan, K. and A. Fabin, 2000. Information Hiding Techniques for Steganography and Digital Watermarking. Artech House, London, UK.

Yeh, C.H. and C.J. Kuo, 1999. Digital watermarking through quasi m-arrays. Proceedings of the IEEE Workshop on Signal Processing Systems, OCt. 20-22, Taipei, Taiwan, pp: 456-461.